Journal of Industrial and Systems Engineering
Vol. 15, No. 2, pp. 59-75
Spring (April) 2023

Polynomial Kernel data-driven robust optimization for modeling competitive pricing problem under influencers marketing

Atiye Yousefi[1], Mir Saman Pishvaee[1*], Babak Amiri[1]

[1]*School of Industrial Engineering, Iran University of Science and Technology, Tehran, Iran*

atyousefi@ymail.com, pishvaee@iust.ac.ir, babakamiri@ iust.ac.ir

Abstract

Empirical studies have indicated that linking advertising and pricing will bring significant advantages to the supply chain components. With the exponential extension of online social networks and society's greater interest in receiving information from this space, many firms have been encouraged to use online social networks and maximize the effects of advertising campaigns; however, literature on designing this type of advertising and linking it with pricing in the supply chain is still rare. To fill this gap, this paper uses a data-driven support vector optimization framework to link influencer-based advertising and pricing in a two-echelon SC. Also, the impact of the passage of time and uncertainty on advertising message diffusion has been examined. The results show that advertising in social media is a complex task and is affected by various factors, such as the time of serving the primary and supporting ads. Based on our results, only after six weeks of releasing the primary ads did the effect of the advertisement decrease significantly. It seems that disseminating supporting advertising messages in advertising campaigns is vital. Also, results obtained from the data-driven robust optimization models show that the slightest change in the degree of conservatism significantly changes the profitability of the company (an increase of only 5% of the degree of conservatism increases profitability by about 1.4 on average), therefore, determining this coefficient has a significant effect on the performance advertising campaigns.

Keywords: Dynamic competitive pricing, time-sensitive advertising, polynomial Kernel, support vector machine, bi-level programming, data-driven programming

1-Introduction

Empirical studies have shown that linking advertising with pricing is quite profitable for the supply chain (SC) components (Sahraeian & Mohagheghian, 2017; Alirezaee & Sadjadi, 2020). This issue has caused several recent studies to address the linking of pricing and advertising decisions in the SC. The research linking advertising with pricing decisions in the supply chain can be divided into two separate lines of work. The first deals with cooperative advertising between supply chain components, while the second relates to non-cooperative advertising between supply chain components, which the current work falls into. A review of this literature reveals that although studies on linking pricing and advertising decisions in the SC are numerous, none has yet addressed advertising in social media marketing, even though with the expansion of customers' activities in social networks and society's interest in obtaining information from this space, electronic businesses worldwide have included using social media to maximize the effects of advertising campaigns.

*Corresponding author
ISSN: 1735-8272,

Influencer marketing is one of the most successful methods in social media advertising campaigns that have recently attracted attention (Vrontis et al., 2021). In this method, some users are selected to diffusion ads in social media because of their experience or charismatic personality to stimulate emotions (Hosseini Bamakan et al., 2019; Zhuang et al., 2021). Such users can reach thousands of people through the content they share (Amiri, et al., 2021), and more importantly, they can guide the emotions and thoughts of the community (Kay et al., 2020). Despite the high efficiency of this method in the success of advertising campaigns, no article has discussed the combination of pricing and advertising decisions, focusing on advertising based on Influencer marketing.

Hence, the mathematical optimization model proposed in this study simultaneously considers the competitive pricing problem and influencer marketing in a two-echelon SC. Our studies show that although practically needed, the literature lacks any effort to design a data-driven framework for social media-based ads considering the effect of time and the probabilistic nature of the information dissemination and examining their impact on pricing in the SC.

Accordingly, the current research is aimed to design a robust data-driven mathematical optimization model capable of modelling the SC competitive pricing under uncertain, time-sensitive social network advertising. The main contributions of this study are as follows:

- A hybrid framework of machine learning and game theory methods has been proposed to model competitive pricing problems under social media influencers' advertising;
- The probabilistic nature of the information dissemination Is based on the polynomial kernel data-driven clustering;
- Advertising costs based on the characteristics of influencers are Optimized;
- The effects of time passing on reducing the influence of advertising messages are analyzed.

The remainder of the paper is organized as follows. Section 2 briefly reviews recent papers that studied the SC's pricing and advertising decisions in the SC and maximizing the influence of online social networks. The problem statement and its assumptions are explained in section 3. The mathematical programming is addressed in section 4, and section 5 describes a data-driven robust optimization model. The solution method is discussed in section 6. Section 7 describes the proposed model by implementing it on the Instagram platform dataset. Finally, section 8 deals with conclusions and gives suggestions for future studies.

2- Literature review

The studies related to this article can be examined in 3 general sections. In the field of non-cooperative advertising between supply chain components recently, Kazemi and Saeed-Mohammadi (2016) coordinated a two-level supply chain consisting of a manufacturer and a retailer in the field of non-cooperative advertising between supply chain components. In this article, two game theory-based models consisting of the manufacturer-Stackelberg and the retailer-Stackelberg game are destined to show the relationship between manufacturer and retailer. Gupta et al. (2019) designed a decentralized three-echelon supply chain consisting of a supplier, a manufacturer, and a retailer. This paper investigates the effect of the market power structure, advertising, and pricing decisions. Also, the uncertainty related to customer demand and final production costs is considered fuzzy variables. Yang et al. (2019) considered a two-echelon supply chain consisting of a supplier and a retailer whose demand is sensitive to the retailer's retail price and advertising expenses. In this paper, the retailer invests part or all of the initial capital exclusively in advertising at the beginning of the sales season. Finally, Geranmayeh and Zaccour (2021) investigated a multi-period supply chain to investigate pricing and advertising decisions. In this paper, the retailer determines the retail prices during the different periods of the selling season, along with the advertising budget, while the manufacturer fixes the wholesale price and its share in the retailer’s advertising cost.

Regarding influence maximization in dynamic social networks, Han et al. (2017) proposed a dynamic framework that examined the main network to identify the subnets by a greedy algorithm that searched in disturbed time intervals. Kuhnle et al. (2018) modelled the network overlapping and heterogeneous dissemination of information by developing an approximate algorithm that considered non-overlapping subnets to maximize the dissemination in the entire network. The proposed framework enables better identification of the subnets in dynamic networks. For considering information dissemination under time and cost constraints, Hu et al. (2018) designed a framework enabling information dissemination in dynamic networks to consider information dissemination under time and cost constraints. This framework shows that user behavior patterns are the main reason for the changing pattern of information dissemination over time. Quan et al. (2018) presented a dynamic framework to predict the re-dissemination of information in social networks. The results indicate that predicting the re-release of information requires influencers' daily and weekly patterns of information diffusion and the power of other users. Min et al. (2020), by considering the time-sensitivity of information dissemination, designed a framework to identify the behavioral pattern of users in information dissemination. The mentioned framework is designed in such a way it can identify the online behavioral patterns of users based on their historical data. Singh & Kailasam (2021), to determine the changing structure of social networks over time (link connection or disconnection), applied the Boltzmann Machine-based deep learning technique to predict the links that may appear in an evolutionary network considering its temporal and structural pattern. The main goal of this paper is to design a framework to influence maximization in an over-time evolving social network. Finally, Baojun et al. (2022) designed an algorithm that identified a set of candidate seeds and then determined the increased influence produced by each node by combining the sub-modularity features to improve the algorithm efficiency for the next period. This paper aims to improve the classical independence cascade dissemination model and consider local topological relations among effective nodes/links as the inter-node influence probability in dynamic social networks.

In the field of mathematical models designed to handle the probabilistic nature of information dissemination, for the first time, He and Kempe (2014) presented an algorithm that estimated the uncertain parameters effective on influence maximization. Next, Chen et al. (2016) used robust optimization concepts to show that uncertainties had highly adverse effects on finding the best possible set of seeds for influence maximization. The above two studies are the most important studies published to show the potential dissemination of information in social networks. Still, none of them proposed any mathematical programming. To develop a mathematical model to control uncertainties of information diffusion in social networks using robustness concepts, Agha Mohammad Ali Kermani et al. (2021) developed a robust optimization model that simultaneously minimized the seed nodes' size and maximized information dissemination. This paper also considers the nodes' heterogeneity in information diffusion.

The above review shows that some areas, such as those in table 1, have not been investigated in the literature. Therefore, the current work proposes a framework that considers social media-based advertising by considering information dissemination's timing and probabilistic nature. It also examines its effects on pricing in a two-tier SC. Accordingly, the present research aims to design a robust data-driven mathematical model for modelling SC competitive pricing under time-sensitive and uncertain social network advertising.

Table 1. The existing research findings

Field	Research Gap
Non-cooperative advertising between supply chain components	Although social media has become the leading platform for corporate advertising, no study has yet addressed link advertising and supply chain pricing based on social media advertising.
The probabilistic nature of information dissemination	Data-driven frameworks have not been used to model and optimize the diffusion of advertising messages.
Influence maximization in dynamic social networks	Uncertainty and time, which are two factors affecting the dissemination of information, have not been sufficiently investigated.

3- Problem statement

From a practical point of view, it is assumed that there is a network involving an online retailer and several suppliers and end customers, where the retailer buys products from suppliers to satisfy the demand of their customers. In this problem, the conflict between suppliers and retailers is modelled as a Stackelberg game.

On the other hand, to advertise on social media platforms and attract customers, the online retailer plans to choose several influencer users by paying an amount in the form of advertising fees to take advantage of these people for presenting the ads. As a platform for advertisements, a network of users of social media is considered, which is represented by a directed graph $\boldsymbol{G}$ ($\boldsymbol{N}$; $\boldsymbol{E}$) (a directed graph), where $\boldsymbol{N}$ (set of vertices) shows the users created by sharing $\boldsymbol{W}$ (set of influencers) and $\boldsymbol{U}$ (set of user followers). $\boldsymbol{E}$ shows the inter-user communication in the network. Information dissemination in the assumed social network is time-dependent and uncertain; in other words, influencers disseminate the retailer's intended advertising messages and send them to their network users at time $\boldsymbol{t = 0}$, but the messages gradually lose their efficacy over time. Another issue this model addresses is the uncertain effect of the message dissemination on the target society, meaning that the extent to which users notice the influencers' messages is uncertain in any period.

To address the above issue, a robust data-based optimization model has been designed to model competitive pricing in a two-echelon SC under uncertain, time-sensitive social network advertising. The designed mathematical programming model can optimize pricing and advertising decisions in a supply chain based on social media by considering the probabilistic nature of the information dissemination and evaluating the effects of time on reducing the influence of advertising messages. In this article, for the first time, the support vector machine-based polynomial kernel is addressed to develop a data-driven robust optimization.

The following are the assumptions of this model:

- The planning horizon is 4 months, containing eight two-week periods;
- Costs paid to the influencers are calculated based on the diffusion of each advertising message, and this cost is assumed to depend directly on the influence network of each influencer. During the first 2 periods (first 4 weeks), influencers have to publish 4 advertising posts every two weeks in the form of Posts, Videos and Stories. It is assumed that the advertising cost is $2000 and the total advertisement budget is $15000;
- The proposed framework considers 5 suppliers and one retailer;
- The parameters related to the rate of attention paid to the message of influencer by users is uncertain, and it is handled based on a data-driven robust optimization model (see section 4);
- The suppliers' market is not independent; that is, decreasing the price in one market causes the demand of another market to decline. In other words, customers move in different markets;
- Instagram social media has been used as a platform for publishing advertising messages;
- When the retailer selects an influencer for advertising, the selected influencer accepts the information. It resends the received message to all its neighboring nodes (all its followers in the social network);
- The reduction factor of advertisement effect on demand in period t is determined based on experts' opinions.

4- Mathematical programming

Based on what was stated, this study aims to develop a data-driven optimization framework to model competitive pricing problems under social media influencers' advertising, which simultaneously considers the probabilistic nature of the information dissemination based on the polynomial kernel data-driven clustering and analyses the effects of time passing on reducing the influence of advertising messages. The variables and parameters used in this model are presented in table 2.

Table 2. The notations used in the mathematical models

Symbol	Description	
Indices		
i	Influencers	$i = 1. \ldots . I$
u	Network users	$u = 1. \ldots . U$
s	Suppliers	$s = 1. \ldots . S$
t	Time period	$t = 1. \ldots . T$
Retailer parameters		
B_t	The budget allocated for advertising in time period $\boldsymbol{t}$	
d_t	Total customer demand in time period $\boldsymbol{t}$	
α_t	Price elasticity coefficient for the retailer's demand function in time period $\boldsymbol{t}$	
β_t	Reduction factor of advertisement effect on demand in time period $\boldsymbol{t}$	
ic_t	The costs paid by influencers for advertising in time period $\boldsymbol{t}$	
$\widetilde{ip}_{iut}$	Rate of attention paid to the message of influencer $\boldsymbol{i}$ by user $\boldsymbol{u}$ in time period $\boldsymbol{t}$ (uncertain parameter)	
N_t	Maximum number of influencers selected for advertising in time period $\boldsymbol{t}$	
M_i	Network influencers whose activation probability is more than a certain limit	
iNe_{it}	The network under the influence of influencer $\boldsymbol{i}$ in time period $\boldsymbol{t}$	
Retailer decision variables		
P_t	The unit price of the product sold to the final customer in time period $\boldsymbol{t}$	
P'_{st}	The unit price of the product purchased from supplier $\boldsymbol{s}$ in time period $\boldsymbol{t}$	
Z_{it}	$\begin{cases} 1 & \text{If micro-influencer } \boldsymbol{i} \text{ is selected in time period } \boldsymbol{t} \\ 0 & \text{Otherwise} \end{cases}$	
li_{iut}	$\begin{cases} 1 & \text{If the advertising message of micro-influencer } \boldsymbol{i} \text{ is noticed by user } \boldsymbol{u} \text{ in time period } \boldsymbol{t} \text{ (link } \boldsymbol{li_{iut}} \text{ is activated)} \\ 0 & \text{Otherwise} \end{cases}$	
Suppliers' parameters		
δ_{st}	Elasticity coefficient of the insider price for supplier $\boldsymbol{s}$ demand function in time period $\boldsymbol{t}$	
γ_{st}	Elasticity coefficient of the rival price for supplier $\boldsymbol{s}$ demand function in time period $\boldsymbol{t}$	
Suppliers' decision variables		
R_{st}	Percent products sold by supplier $\boldsymbol{s}$ to the retailer in time period $\boldsymbol{t}$	
P'_{st}	The unit price of the product sold by supplier $\boldsymbol{s}$ to the retailer in time period $\boldsymbol{t}$	
P''_{st}	The unit price of the product sold by supplier $\boldsymbol{s}$ to the final customer in time period $\boldsymbol{t}$ (without the retailer intermediary)	

Considering the definition of the problem, parameters and variables, the bi-level mathematical model for competitive pricing under uncertain, time-sensitive social network advertising is as follows:

$$\max \pi_{tr} = \sum_{t=1}^{T} d_t . P_t - \sum_{t=1}^{T} \sum_{s=1}^{S} P'_{st} . R_{st} - \sum_{t=1}^{T} \sum_{i=1}^{I} ic_t \, Z_{it} \tag{1}$$

$$d_t = \beta_t \left[\left(\sum_{t=1}^{T} \sum_{i=1}^{I} \sum_{u=1}^{U} \widetilde{ip}_{iut} \, . \, iNe_{it} \right) li_{iut} \right] - \alpha \sum_{t=1}^{T} P_t \qquad \forall t \tag{2}$$

$$d_t \le \sum_{s=1}^{S} \sum_{t=1}^{T} R_s \tag{2}$$

$$P_t \geq P'_{st} \tag{2}$$

$$ic_t \leq B_t \qquad \forall t \tag{3}$$

$$\sum_{u \in iNe_{it}} li_{iut} \leq M\,.Z_{it} \qquad \forall i.t \tag{4}$$

$$\sum_{u \in N_i} li_{iut} \geq Z_{i(t+1)} - Z_{it} \qquad \forall i.t \tag{5}$$

$$M.Z_{i(t+1)} \geq \sum_{u \in N_i} li_{iut} \qquad \forall g.t \tag{6}$$

$$Z_{it} \leq Z_{i(t+1)} \qquad \forall g.t \tag{7}$$

$$\sum_{u \in iNe_{it}} li_{iu(t+1)} \leq M.\left[Z_{i(t+1)} - Z_{it}\right] \qquad \forall i.t \tag{8}$$

$$\sum_{i=1}^{I} Z_{it} \leq K_t \qquad \forall t \tag{9}$$

$$P'_{st}\,.P_t \geq 0 \qquad \forall s.t \tag{10}$$

$$Z_{it} \in \{0\,.1\} \qquad \forall g.a.i.t \tag{11}$$

$$max\ \pi_{st} = \sum_{t=1}^{T}\sum_{s=1}^{S} P'_{st}R_{st} + \sum_{t=1}^{T}\sum_{s=1}^{S} P''_{st}\,(1 - R_{st}) \tag{12}$$

$$R_{st} = d_t - \delta_{st}\,P'_{st} + \gamma_{st}\,P'_{\neg st} \qquad \forall s.t \tag{13}$$

$$R_{st}.P'_{st}.P''_{st} \geq 0 \qquad \forall s.t \tag{14}$$

Equation (1) represents the retailer’s profit objective function, which is designed in a way that, in addition to showing the retailer's profit from customer demand, it also selects the minimum influencers based on the overall advertising budget. Equation (2) shows the retailer's demand function. In this equation, in addition to showing the effect of price on the reduction of demand, the retailer's profit due to the advertisements seen by the users is also considered. To show the reduction factor of advertisement effect on demand in time period $\boldsymbol{t}$, the $\boldsymbol{\beta_t}$ has been used. Also, The rate of attention paid to the message of influencer $\boldsymbol{i}$ by user $\boldsymbol{u}$ in time period $\boldsymbol{t}$ is also considered an uncertain parameter. Equation (3) shows the retailer’s maximum budget considered for advertising. Equation (4) refers to the fact that if the influencers are selected as advertisement providers, they will send the message to their network members. Based on this constraint, if a node is activated, its leaving lines can be active or inactive. The classic information diffusion models have a probabilistic constraint, which indicates that each of its leaving lines is activated simultaneously with a certain probability when a node is activated. To model this assumption, use is made of the idea of the linear threshold model that states if the activation probability of a leaving wing exceeds a certain limit, it will be activated. Therefore, this study has assumed that if a person is activated (receives the advertising message), those nodes of his network will be activated to which the probability of sending a message is higher than a certain limit. To model the probability nature of ads diffusion in social networks, we used index $\boldsymbol{M_i}$, which refers to those network influencers whose activation probability is more than a certain limit, which is the limit of the communication function of each user. Based on these three indices, equation (5) concludes that if a node is activated at time $\boldsymbol{t+1}$, at least one of its entering lines has been active at time $\boldsymbol{t}$, and if a node is active at both times $\boldsymbol{t}$ and $\boldsymbol{t+1}$, its entering lines can be active or inactive at time $\boldsymbol{t}$. Constraint (6) ensures that the model progressiveness assumption is satisfied. Constraints (7) and (8) allow each node’s leaving lines to be activated only once. Finally, constraint (9) shows the maximum number of the detected influencers. To show the competitive retailer-supplier relationship to determine the purchase price of products from suppliers, the first term of the suppliers' objective function (equation 12) calculates the price of the percent product sold by supplier $\boldsymbol{s}$ to the retailer. The second term of equation (12) also finds the profit from unit product sold by supplier $\boldsymbol{s}$ to the final customer. Equation (13) calculates the percent products sold by supplier $\boldsymbol{s}$ to the retailer, $\boldsymbol{P'_s}$ is the price of unit product provided by supplier$\boldsymbol{s}$, and $\boldsymbol{P'_{\neg s}}$ is the price of unit raw material provided by the competitors of supplier$\boldsymbol{s}$.

5- Data-driven robust optimization model

Robust optimization is a field of mathematical optimization theory that deals with optimization problems in which a certain measure of robustness is sought against an uncertainty that can be represented as deterministic variability in the value of the problem's parameters and its solution (Ben-Tal et al., 2009). Among several general approaches developed for robust optimization, one is "robust programming based on closed convex uncertainty sets". The literature developed in this field has focused on the design of convex uncertainty sets, which, according to their nature, try to control the solutions' degree of conservatism. The following linear programming describes the framework of these problems:

$$\min_{x \in X} C^T X$$
$$S.t. \quad \sum_i \tilde{a}_i x_i \leq b \tag{15}$$

Where $\tilde{a}_i$ coefficients are uncertain. To design the equivalent robust programming problem, each uncertain $\tilde{a}_i$ coefficient is modelled as an independent, symmetric, bounded random variable similar to equation (16), where a_i is the nominal value and $\hat{a}_i$ is the range of variations.

$$\tilde{a}_i = a_i + \xi_i \hat{a}_i \quad \forall i \tag{16}$$

Based on the above definition, the robust counterpart of the main constraint can be shown as follows:

$$\sum_i a_i x_i + \max_{\xi \in \mathcal{U}} \sum_i \xi_i \hat{a}_i x_i \leq b \tag{17}$$

Where ξ_i is symmetrically distributed in the [-1, 1] interval and belongs to a predefined uncertainty set $\mathcal{U}$.

Although several studies have the uncertainty set U, such as box uncertainty set, box-elliptic and box-elliptic-polyhedron uncertainty set, a common drawback in all of them is that the design of their structures lacks the information of uncertain parameters. Under such conditions, finding a minimum-size uncertainty set that covers all the possible realizations of uncertain parameters is almost impossible. In other words, robust optimization approaches arising from the uncertainty set are based on the assumption that there is no knowledge of the uncertain parameter values. Although it may not be possible to estimate the distribution function in real cases due to the lack of rich historical data, the empirical data can be used to define a given uncertainty set. So, in recent years, robust data-driven optimization approaches have been proposed to extract an uncertainty set from uncertain parameter information (Hodjat Panah & Rasouli, 2021; Mohammadi et al., 2021).

Nevertheless, formulating an appropriate uncertainty set for modelling robust optimization problems remains challenging. An optimal uncertainty set should adapt flexibly to the intrinsic structure behind the data. From a machine learning perspective, constructing uncertainty sets based on historical data can be considered an unsupervised learning problem. Several effective unsupervised learning models, for example, kernel density estimation (KDE) and support vector machines (SVM), can provide powerful representations of data distributions. In principle, such machine learning tools can estimate data density with sufficient accuracy.

Among approaches developed in recent years to use data to design uncertainty sets more precisely, the one proposed by Shang et al. (2017) uses the support vector machine (SVM) to make an uncertainty set. The Support Vector Machine is a supervised learning algorithm mostly used for classification. The main idea is that based on the labeled data (training data), the algorithm tries to find the optimal hyperplane which can be used to classify new data points. In two dimensions, the hyperplane is a simple line (Pisner and Schnyer, 2020). SVM algorithms use a set of mathematical functions that are defined as the kernel. The function of kernel is to take data as input and transform it into the required form. Different SVM algorithms use different types of kernel functions. These functions can be different types. For example, linear, nonlinear, polynomial, radial basis function (RBF), and sigmoid (Liu et al., 2014).

When selecting a kernel function for an industrial classification problem, there is no one-size-fits-all answer. It depends on the data characteristics, such as size, dimensionality, distribution, noise, and outliers—generally, a polynomial kernel if it has nonlinear patterns or interactions between features.

Due to the nonlinear patterns of uncertain parameter of this paper (Rate of attention paid to the influencer message of influencer *i* by user *u* in period *t*, we used the Polynomial Kernel method to take data as input and transform it into the required form.

To explain this model, let set $D = \{s_k\}_{k=1}^{N}$ contains N data samples. The main SVC algorithm seeks a closed hypersphere with minimum volume to accommodate almost all data samples. So, the mathematical programming based on the SVM model can be formulated as follows:

$$\begin{aligned} &\min R^2 \\ &s.t. \quad \|\emptyset(s_k) - c\|^2 \le R^2 \qquad \forall k = 1.2.\dots.N \end{aligned} \tag{18}$$

Where $\emptyset(s_k)$ is a transformation function, c is the center, and R is the radius of the hypersphere. To create flexibility in the cluster boundaries and prevent being influenced by outlier data, the variableξ, larger than 0, is added to the right side of the constraints. This constraint makes it possible that outlier data are not included in the hypersphere.

$$\|\emptyset(s_k) - c\|^2 \le R^2 + \xi_k \qquad \forall k = 1, 2, \dots, N \tag{19}$$

By writing the Lagrange equation considering β_k and α_k as Lagrangian coefficients, we will have:

$$L(R, c, \lambda) = R^2 - \sum_k \beta_k \left[R^2 + \xi_k - \|\emptyset(s_k) - c\|^2\right] - \sum_k \alpha_k \xi_k + P \sum_k \xi_k \tag{20}$$

WP is the penalty coefficient applied if noise points are activated. By deriving L concerning R and c we will have:

$$\frac{\partial L}{\partial R} = 0 \quad \rightarrow \quad \sum_k \lambda_k = 1 \tag{21}$$

$$\frac{\partial L}{\partial c} = 0 \quad \rightarrow \quad c = \sum_k \lambda_k \, \emptyset(s_k) \tag{22}$$

Considering KKT supplementary conditions, we will also have the following:

$$\alpha_k \xi_k = 0 \tag{23}$$

$$\beta_k \left[R^2 + \xi_k - \|\emptyset(s_k) - c\|^2\right] = 0 \tag{24}$$

Considering the above equation, we will have:

$$\|\emptyset(s_k) - c\|^2 \le R^2 \quad \rightarrow \quad \lambda_k = 0 \ \ and\ k\ is\ a\ inherent\ point\ (ID) \tag{25}$$

$$\|\emptyset(s_k) - c\|^2 = R^2 \quad \rightarrow \quad 0 < \lambda_k < P \ \ and\ k\ is\ a\ SV\ point \tag{26}$$

$$\|\emptyset(s_k) - c\|^2 > R^2 \quad \rightarrow \quad \lambda_k = P \quad and\ k\ is\ a\ BSV\ point \tag{27}$$

According to equations (25) to (27), if the borders of the sphere and point k neighborhood are a distance apart, then $\lambda_k = 0$ for that point, it lies inside the sphere, and it cannot be used to obtain the support vectors, but if the two borders touch, $\lambda_k > 0$ for that point and such points are called support vectors. Using equations (23) and (24), Lagrange equation (20) will turn into the following quadratic programming model:

$$\max \sum_{k=1}^{N} \lambda_k K(s_k, s_k) - \sum_{k=1}^{N} \sum_{k\prime=1}^{N} \lambda_k \lambda_{k\prime} K(s_k, s_{k\prime})$$
$$\sum_{k} \lambda_k = 1 \tag{28}$$
$$\lambda_k \geq 0 \qquad \forall k = 1.2. \dots . N$$

Where $K(s_k, s_{k\prime})$ is the kernel function, the determination of which necessitates determining the kernel type first. By considering the polynomial kernel method, we will have the following (Vinge & McKelvey, 2019):

$$K(s_k, s_{k\prime}) = \emptyset^T(s_k)\emptyset(s_{k\prime}) \xrightarrow{By\ Considering\ Polynomial\ Kernel} K(s_k, s_{k\prime}) = ({s_k}^T s_{k\prime} + 1)^2 \tag{31}$$

Using the polynomial theorem, we will have

$$\left(\sum_{k=1}^{N} {s_k}^T {s'}_k + 1\right)^2 = \sum_{k=1}^{N} (s_k^2)\,({s'}_k^2) + \sum_{k=2}^{N} \sum_{k'=1}^{k-1} (\sqrt{2} s_k s_{k\prime})\,(\sqrt{2} s'_k s'_{k\prime})$$
$$+ \sum_{k=1}^{N} (\sqrt{2} s_k)\,(\sqrt{2} {s'}_k) + 1 \tag{32}$$

To calculate the kernel, we can consider the matrix K as follows:

$$K = \left\{k(s_k s_{k'})\right\}_{(k,k'=1)}^{N} \tag{34}$$

$$K = \begin{pmatrix} \emptyset^T(s_1)\emptyset(s_1) & \cdots & \emptyset^T(s_1)\emptyset(s_n) \\ \vdots & \ddots & \vdots \\ \emptyset^T(s_n)\emptyset(s_1) & \cdots & \emptyset^T(s_n)\emptyset(s_n) \end{pmatrix} \tag{35}$$

By placing the matrix K in problem (28) and obtaining the optimal amount of λ_k, the radius of the hypersphere is calculated as follows:

$$R^2 = K(s_k, s_{k\prime}) - 2\sum_{k=1}^{N} K(s_k, s_k)\lambda_k + \sum_{k=1}^{N} \sum_{k'=1}^{N} \lambda_k \lambda_{k'} K(s_k, s_{k'}) \qquad k \in SV \tag{36}$$

The $U(s)$ uncertainty set that covers the inner region of the hypersphere (radius R) is defined as follows:

$$U(s) = \left\{ s \,\middle|\, \sum_{k \neq ID} ({s_k}^T s_{k\prime} + 1)^2 \lambda_k \leq \sum_{k \neq ID} ({s_k}^T s_{k\prime} + 1)^2 \lambda_k \qquad k \in SV \right\} \tag{39}$$

Using auxiliary variables $F = [f_1. \dots . f_N]$ instead of $({s_k}^T s_{k\prime} + 1)^2$ and define the $\sum_{k \neq ID} ({s_k}^T s_{k\prime} + 1)^2 \lambda_k = \Phi$, the uncertainty set (40) can be rewritten as follows:

$$U(s) = \left\{ s \quad \begin{array}{c} \exists f_k, k \neq ID, s.t \\ \sum_{k \neq ID} \lambda_k f_k \leq \Phi \\ \sum_{k \neq ID} ({s_k}^T s_{k\prime} + 1)^2 \lambda_k = \Phi \end{array} \right\} \tag{40}$$

Under the proposed uncertainty set, the robust counterpart of model (15) is formulated as follows:

$$\begin{aligned} &\min_{x \in X} C^T X \\ &S.t. \quad \max_{a \in U(s)} a^T x \leq b \end{aligned} \tag{41}$$

To obtain the equivalent LP form of (41), the inner optimization problem can be written as follows:

$$\begin{aligned} &\max_{S, f_k} S^T X \\ &S.t. \quad \sum_{k \neq ID} \lambda_k f_k \leq \Phi \\ &\sum_{k \neq ID} \left({s_k}^T s_{k'} + 1\right)^2 \lambda_k = \Phi \end{aligned} \tag{42}$$

By introducing duality variables α_k and β_k, the internal optimization problem (42) can be transformed into its dual forms as (Mohseni and Pishvaee, 2020):

$$\begin{aligned} &\min_{\alpha_k, \beta_k} \sum_{k \neq ID} \alpha_k \Phi + \beta_k \Phi \\ &S.t. \quad \sum_{k \in SV} f_k \alpha_k + \sum_{k \neq ID} ({s_k}^T s_{k\prime} + 1)^2 \beta_k = 0 \\ &\alpha_k, \beta_k \geq 0 \end{aligned} \tag{43}$$

In corporation of the dual problem (48) into the robust optimization problem (46) obtains the following robust counterpart problem:

$$\begin{aligned} &\min_{x \in X} C^T X \\ &S.t. \quad \sum_{k \neq ID} \alpha_k \Phi + \beta_k \Phi \leq b \\ &\sum_{k \in SV} f_k \alpha_k + \sum_{k \neq ID} ({s_k}^T s_{k\prime} + 1)^2 \beta_k = 0 \\ &\alpha_k, \beta_k \geq 0 \end{aligned} \tag{44}$$

6- Solution method

Bilevel optimization is defined as a mathematical program where an optimization problem contains another optimization problem as a constraint. Consequently, bilevel optimization can model hierarchical decision processes (Sinha et al., 2018). Among several methods in the literature to solve bi-level programming (BLP) problems, Karush–Kuhn–Tucker (KKT) conditions try to add the second-level problem to the first-level one under the KKT conditions and be changed to a single-objective problem; if so, it is converted into a nonlinear mix integer programming using 0, 1 variable (Allende and Still, 2013). The optimality conditions for a constrained local optimum are called the Karush Kuhn Tucker (KKT) conditions, and they play an important role in constrained optimization theory and algorithm development. The KKT conditions for optimality are necessary for a solution to be optimal in a mathematical optimization problem. They are necessary and sufficient conditions for a local minimum in nonlinear programming problems (Jahn, 2017). To consider the Karush-Kuhn-Tucker condition for solving bilevel optimization problems, consider the following compact model:

$$\begin{aligned}
&\max_{x_1} F(x_1.x_2)\\
&s.t.\ G\ (x_1.x_2) \leq B\\
&\qquad \max_{x_2} f(x_1.x_2)\\
&\qquad s.t.\ g\ (x_1.x_2) \leq C\\
&\qquad x_1.x_2 0
\end{aligned} \tag{45}$$

Considering a bi-level problem similar to equations (45), the KKT conditions (necessary for optimality) can be expressed in the form of equations ((46)-(49)).

$$\nabla f(x) = \sum_{i-1}^{m} \lambda_i \nabla g_i(x) \qquad i = 1. \dots . m \tag{46}$$

$$\lambda_i g_i(x) = 0 \qquad i = 1. \dots . m \tag{47}$$

$$g_i(x) \leq 0 \qquad i = 1. \dots . m \tag{48}$$

$$\lambda_i \geq 0 \qquad i = 1. \dots . m \tag{49}$$

To use the KKT conditions to convert a bi-level to a single-level problem, equations ((50)-(53)) are added to the latter, causing the bi-level one to be omitted.

$$-\frac{\partial f(x_1.x_2)}{\partial(x_1.x_2)} + \frac{\partial g_i\ (x_1.x_2)}{\partial(x_1.x_2)} = 0 \tag{50}$$

$$(g_i\ (x_1.x_2) - B) * \lambda_i = 0 \tag{51}$$

$$g_i\ (x_1.x_2) - B \leq 0 \tag{52}$$

$$0 \leq \lambda_i \leq 1 \tag{53}$$

By multiplying the Lagrange coefficients by the constraints (similar to equation (51)), the model is changed to a non-linear problem, which becomes linear by binary variable $\boldsymbol{\omega_i}$ and the following auxiliary constraints:

$$\lambda_i \leq M * \omega_i \qquad \forall\ i \tag{54}$$

$$-g_i(x) \leq M(1 - \omega_i) \qquad \forall\ i \tag{55}$$

Hence, considering the KKT conditions, the compact model (45) will be rewritten as follows:

$$\begin{aligned}
&Max F(x_1.x_2)\\
&s.t.\\
&G\ (x_1.x_2) \leq B\\
&g\ (x_1.x_2) \leq B\\
&-\frac{\partial f(x_1.x_2)}{\partial(x_1.x_2)} + \frac{\partial g\ (x_1.x_2)}{\partial(x_1.x_2)} = 0\\
&\lambda \leq M * \omega\\
&-g\ (x_1.x_2) - B \leq M(1 - \omega)\\
&g\ (x_1.x_2) - B \leq 0\\
&0 \leq \lambda \leq 1
\end{aligned} \tag{54}$$

7- Case study

To demonstrate the efficiency of the developed models, some information in the dataset introduced by Kim et al. (2021) has been used. This dataset contains a lot of information, including hashtags used by influencers, No. Of likes, comments, extent of influence network, etc., of 33935 Instagram influencers. Also, all the results have been calculated by the CPLEX optimization software on a Core i7 computer with an 8.0 GB RAM.

As the first analysis, we aimed to find the effects of the passage of time on the success of disseminating advertising messages and on the increased demand. To do this, demand variations versus those in the network of influencers and the coefficient of the reduced effect of advertising on demand are plotted in figure 1 for the 8 periods. As can be seen, with the increase of the network affected by advertising messages, the demand increases with a rapid trend, and the third period, two periods after the start of advertising, reaches its maximum level, but after that, it faces a sharp fall. Based on these analyses, companies can republish advertising messages with appropriate timing to prevent the advertising effects from dropping.

In this study, we used the experts' opinions to determine the reduction factor of advertising effects on demand. However, choosing the exact time of the re-dissemination of advertising messages can be predicted based on historical data. Based on the above discussions, it can be acknowledged that advertising in social media is very complicated, the efficacy of which depends on several factors, such as the proper dissemination time by the right people.

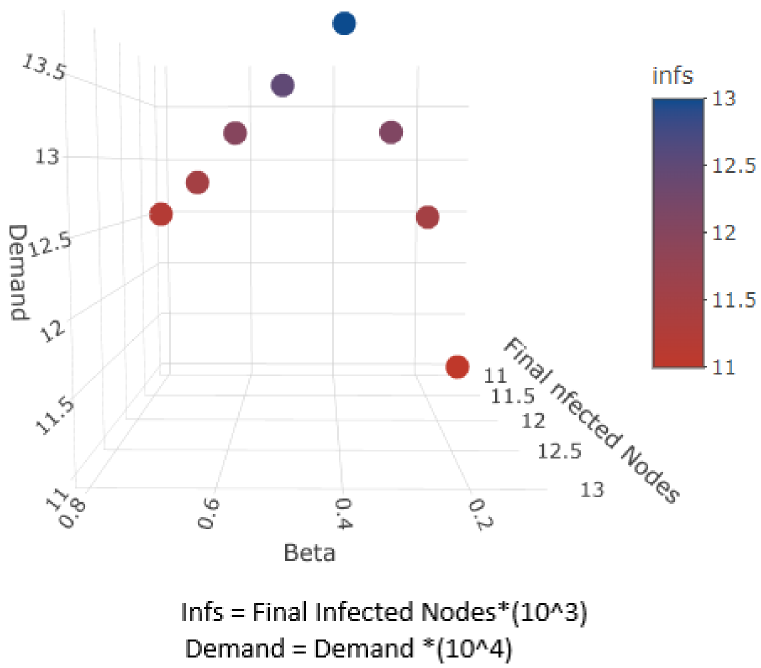

Fig 1. Effects of time on the dissemination of advertising messages

To measure the performance of SVC-based uncertainty sets, this section compares the uncertainty set formed by SVC to the classical box and polynomial uncertainty sets. To this end, we used the Monte Carlo Simulation (MCS) to generate random realizations to analyze the quality of the solutions found by SVCU (support vector clustering uncertainty), PU (Polynomial uncertainty) and BU (box uncertainty).

Figure 2 shows the conservatism level of robust solutions, and retailer profit is shown for different values of μ (μ is a coefficient that controls the distance of random realizations from data points). The conservatism level of robust solutions indicates the tendency of a robust model to produce a solution that avoids violating the constraints. In other words, the more a model avoids violation of constraints, the more conservative it is (Lan, 2021).

At a similar conservatism level, the value of the objective function of the SVCU model is more than that of the BU and PU robust optimization models, which means the SVCU make the same-quality solutions in terms of robustness with lower costs. Another point worth considering is SVCU merit in creating higher conservatism-level robust solutions that are impossible to achieve by other robust optimization models, especially when the conservatism level increases (Mohseni and Pishvaee, 2020).

Also, the gap between the data-driven robust optimization, BU and PU model is narrowed as the value grows. This is because, for small values, the distribution of the random realizations is almost identical to that of the data points used to construct the uncertainty sets. Still, when the value increases, the random realizations diverge from the data points and lie outside the uncertainty set more than before.

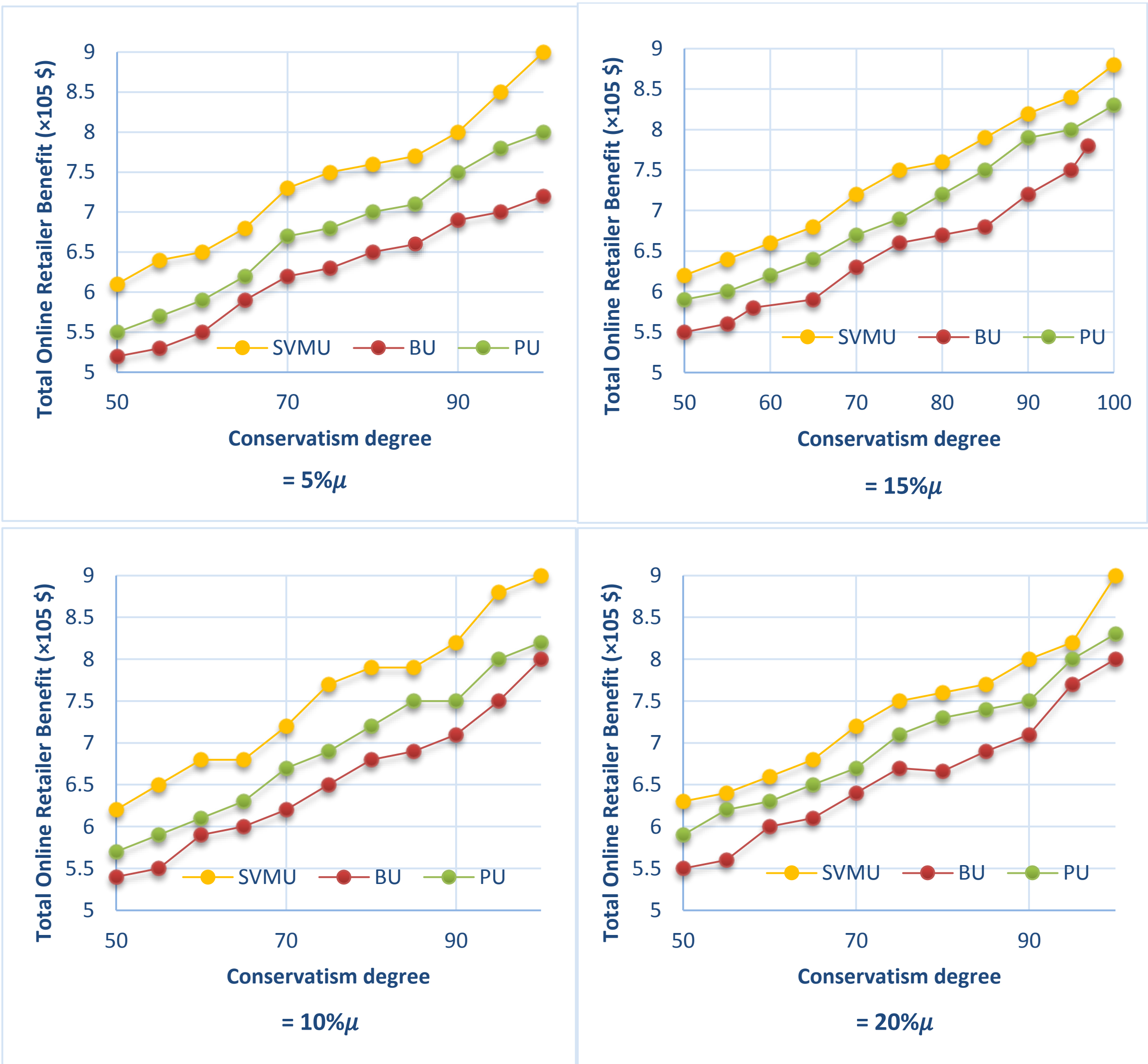

Fig 2. Performance of uncertainty sets

Based on this, it can be concluded that the proposed data-driven robust optimization model is more efficient than traditional models in conservatism-level cost management.

As the last presented analysis, the price sensitivity effect on retailer and supplier profit has been investigated. As seen in figure 3, by increasing the parameter α (price elasticity coefficient for the retailer's demand function), the sensitivity of demand to price changes increases, and the demand for a specific price decrease. With the increase of the price elasticity coefficient, which refers to the market's competitive nature, the average price decreases at all levels of the supply chain, reducing the profit of all chain members, especially retailers. This happens for products in a market with perfect competition where the chain members are forced to reduce prices for more profit. Based on the above analysis, small price changes in competitive markets will significantly impact these markets, so businesses should decide more carefully about price determination.

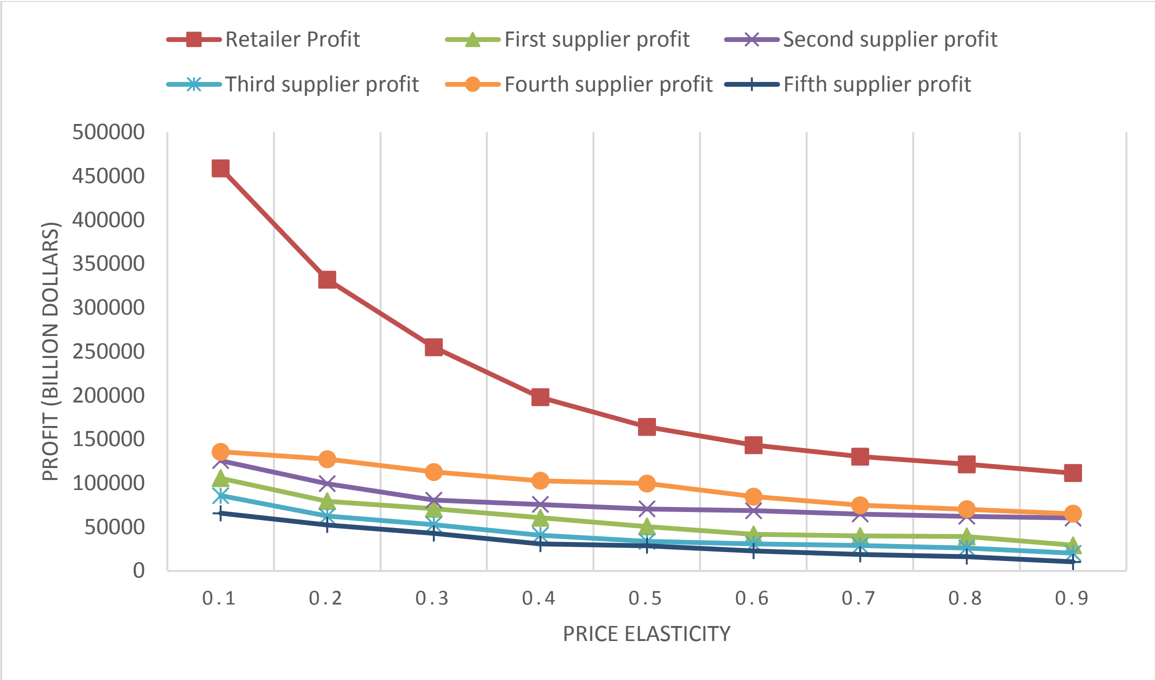

Fig 3. The effect of price sensitivity on retailer and supplier profit

7-1- Managerial insights

Based on the above analyses, generally, it can be concluded that companies face 4 significant challenges in advertising diffusion using influencers in social media that need to be addressed:
Based on the above analyses, generally, it can be concluded that companies face 4 major challenges in advertising diffusion using influencers in social media that need to be addressed:

1) With time, the effectiveness of social network ads decreases significantly;
2) Releasing successful advertising is more than choosing the right influencers. This issue is affected by the production of attractive content and publishing it at the right time;
3) By starting advertising, the activation speed of celebrities with a huge communication network is much higher than those of other influencers, but it decreases over time (see Fig. 2);
4) Due to the probabilistic nature of social media advertising, using the appropriate uncertainty-based optimization models is necessary to provide stability.

In order to deal with the above challenges, it is suggested that 1) the company use historical data to determine the exact time of the re-dissemination of advertising messages to prevent the loss of ads effect, 2) selecting the influencers should be done more carefully, and attention should be paid to the more important features than their huge communication network, 3) in order to get more realistic results, companies should use robust optimization models to deal with the uncertainty of advertising messages, but determining the degree of conservatism is an important decision that has a significant impact on cost controlling.

8- Conclusions and future studies

This research has considered a data-driven optimization model and influencer-based advertisements for pricing in a two-echelon supply chain by studying the uncertain dissemination of advertising messages and the effects of time on the reduced effects of these advertisements. The proposed approach can be effective in several research fields and provide a basis to develop the related literature by 1) presenting a new approach for data-driven pricing in supply chains, 2) creating a new direction to link pricing decisions and advertising in supply chains through social media-based advertising, 3) considering time-sensitive social network advertising and 4) using a polynomial kernel data-driven robust model to handle the non-robustness of decisions in facing data variations.

The analysis results obtained from this approach show that this study can continue in the following research directions: 1) Examining the historical data on the reduced effects of advertising messages and estimating the related collapse function aiming at injecting advertising messages in proper time periods, 2) Identifying the features of appealing ads for the audience and designing them attractively, 3) Using the behavioral sciences (e.g., nudge) to enhance the advertisement efficacy and 4) Developing the uncertain dissemination of advertising messages using data-driven adaptive models.

References

Agha Mohammad Ali Kermani, M., Ghesmati, R., Pishvaee, M.S. (2021). A robust optimization model for influence maximization in social networks with heterogeneous nodes. *IEEE Computer Society*, 8(1), 17.

Alirezaee, A., Sadjadi, S. J. (2020). A game theoretic approach to Pricing and Cooperative advertising in a multi-retailer supply chain. *Journal of Industrial and Systems Engineering*, 12(4), 154-171.

Allende, G.B., Still, G. (2013). Solving bilevel programs with the KKT-approach. *Mathematical Programming*. 138(1-2), 309–332.

Amiri, B., Fathian, M., Asaadi, E. (2021). Influence maximization in complex social networks based on community structure. *Journal of Industrial and Systems Engineering*, 13(3), 16-40.

Baojun, F., Zhang, J., Bai, H., Yang, Y., He, Y. (2022). An Influence Maximization Algorithm for Dynamic Social Networks Based on Effective Links. *Entropy*, 24(7). 904.

Ben-Tal, A., El Ghaoui, L., Nemirovski, A. (2009). *Robust optimization*. Princeton university press, 28.

Chen, W., Lin, T., Tan, Z., Zhao, M., Zhou, X. (2016). *Robust Influence Maximization*. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16). Association for Computing Machinery, New York, NY, USA, 795–804.

Domingos, P., Richardson, M. (2001). *Mining the network value of customers*. In: Proceedings of the 7th ACM SIGKDD international conference on Knowledge discovery and data mining. Association for Computing Machinery, New York, NY, USA, 57–66.

Geranmayeh, A., Zaccour, G. (2021). Pricing and advertising in a supply chain in the presence of strategic consumers. *Omega*, 101, 102239.

Gupta, R., Biswas, I., Kumar, S. (2019). Pricing decisions for three-echelon supply chain with advertising and quality effort-dependent fuzzy demand, International Journal of Production Research, 57(9), 2715-2731.

Han, M., Yan, M., Cai, Z., Li, Y., Cai, X., and Yu, J. (2017). Influence maximization by probing partial communities in dynamic online social networks. *Transactions on Emerging Telecommunications Technologies*, 28: e3054.

He, X., Kempe, D. (2014). *Stability of influence maximization*. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 1256–1265.

HodjatPanah, N., Rasouli, M. R. (2021). A data-driven model for credit scoring of loan applicants within a crowdfunding scenario in a P2P lending platform in Iran. *Journal of Industrial and Systems Engineering*, 13(4), 288-306.

Hosseini Bamakan, S.M., Nurgaliev, I., Qu., Q. (2019). Opinion leader detection: A methodological review. Expert Systems with Applications, 115, 200-222.

Hu, M., Liu, Q., Huang, H., Jia, X. (2018). *Time-Sensitive Influence Maximization in Social Networks*. In: 18th International Conference on Communication Technology (ICCT), Chongqing, China, 1351-1356.

Jahn, J. (2017). Karush–Kuhn–Tucker conditions in set optimization. *Journal of Optimization Theory and Applications*, 172(3), 707-725.

Kay, S., Mulcahy, R., Parkinson, J. (2020). When less is more: the impact of macro and micro social media influencers' disclosure. Journal of Marketing Management, 36(3-4), 248-278.

Kazemi, A., SaeedMohammadi, Z. (2016). A Model for Cooperative Advertising and Pricing Decisions in Manufacturer-Retailer Supply Chain with Discount: A Game Theory Approach. International Journal of Supply and Operations Management, 2(4), 1035-1063.

Kennedy, A.P., Sethi, S.P., Siu, C.C., Yam, S.C.P. (2021). Cooperative Advertising in a Dynamic Three-Echelon Supply Chain. *Production and Operations Management*, 30(11), 3881-3905.

Kim, S., Jiang, J.Y., Wang, W. (2021). *Discovering Undisclosed Paid Partnership on social media via Aspect-Attentive Sponsored Post Learning*. In: Proceedings of Web Search and Data Mining, Association for Computing Machinery, New York, USA.

Kuhnle, A., Alim, M. A., Li, X., Zhang, H., Thai, M. T. (2018*). Multiplex Influence Maximization in Online Social Networks with Heterogeneous Diffusion Models*. In: IEEE Transactions on Computational Social Systems, Lanzhou, China, 5(2), 418-429.

Lan, Y. (2021). Continuous Control of Conservatism for Robust Optimization by Adjustable Regret. *ArXiv (Cornell University)*. https://doi.org/10.48550/arxiv.2105.05536.

Li, M., Zhang, X., Dan, B. (2021). Cooperative advertising and pricing in an O2O supply chain with buy-online-and-pick-up-in-store. *International Transactions in Operational Research*, 28(4), 2033-2054.

Liu, L., Shen, B., Wang, X. (2014). *Research on Kernel Function of Support Vector Machine*. In: Huang, YM., Chao, HC., Deng, DJ., Park, J. (eds) Advanced Technologies, Embedded and Multimedia for Human-centric Computing. Lecture Notes in Electrical Engineering, 260. Springer, Dordrecht.

Min, H., Cao, J., Yuan, T., Liu, B. (2020). Topic based time-sensitive influence maximization in online social networks. *World Wide Web*, 23(3), 1831–1859.

Mohammadi, M., Rasouli, M. R., Pishvaee, M. S. (2021). Data-driven approaches for customer-centric and service-dominant value propositions: A systematic literature review. *Journal of Industrial and Systems Engineering*, 14(1), 51-75.

Mohseni, M., Pishvaee, M.S. (2020). Data-driven robust optimization for wastewater sludge-to-biodiesel supply chain design, *Computers & Industrial Engineering*, 139, 105944.

Mokhlesian, M., Zegordi, S. H. (2018). Pricing and advertising decisions in a dominant-retailer supply chain: A multi-follower bi-level programming approach. *Scientia Iranica*, 25(4), 2254-2266.

Pisner, D. A., Schnyer, D. M. (2020). *Support vector machine*. In: Machine Learning: Methods and Applications to Brain Disorders. Academic Press, 101-121.

Quan, Y., Jia, Y., Zhou, B., Han, W., Li, S. (2018). Repost prediction incorporating time-sensitive mutual influence in social network. *Journal of Computational Science*, 28(9), 217-227.

Sahraeian, R., Mohagheghian, E. (2017). A game Theoretic Approach to Pricing, Advertising and Collection Decisions adjustment in a closed-loop supply chain. *Journal of Industrial and Systems Engineering*, 10(2), 53-74.

Shang, C., Huang, X., You, F. (2017). Data-driven robust optimization based on kernel learning. *Computers & Chemical Engineering*, 106, 464–479.

Sadjadi, S.J., Alirezaee, A. (2020). Impact of pricing structure on supply chain coordination with cooperative advertising. *RAIRO-Operations Research*, 54(6), 1613-1629.

Singh, A., Kailasam, L. (2021). Link Prediction-Based Influence Maximization in Online Social Networks. *Neurocomputing*, 453(C), 151-163.

Sinha, A., Malo, P., Deb, K. (2018). A Review on Bilevel Optimization: From Classical to Evolutionary Approaches and Applications, IEEE Transactions on Evolutionary Computation, 22(2), 276-295.

Vinge, R., McKelvey, T. (2019, September 1). Understanding Support Vector Machines with Polynomial Kernels. *IEEE Xplore*. https://doi.org/10.23919/EUSIPCO.2019.8903042.

Vrontis, D., Makrides, A., Christofi, M., & Thrassou, A. (2021). Social media influencer marketing: A systematic review, integrative framework and future research agenda. *International Journal of Consumer Studies*, 45(4), 617-644.

Xiao, M., Wang, R., Chan-Olmsted, S. (2018). Factors affecting YouTube influencer marketing credibility: a heuristic-systematic model. Journal of Media Business Studies, 15(3), 188-213.

Yang, H., Chu, L., Wan, H. (2019). Advertising and Pricing Policies in a Two-Echelon Supply Chain with a Capital-Constrained Retailer, RAIRO - Operations Research, 53(4), 1331-1342.

Zhuang, H., Sun, Y., Tang, J., Zhang, J., Sun, X. (2013). *Influence Maximization in Dynamic Social Networks*. In: 13th International Conference on Data Mining, Dallas, TX, USA, 1313-1318.